Published in final edited form as:

J Clin Gastroenterol. 2019 January ; 53(1): 40–50. doi:10.1097/MCG.0000000000000872.

Race, Age, and Geography Impact Hepatitis C Genotype Distribution in the United States

Stuart C. Gordon, MD[*], **Sheri Trudeau, MPH**[†], **Jia Li, PhD**[†], **Yueren Zhou, MS**[†], **Loralee B. Rupp, MBA**[‡], **Scott D. Holmberg, PhD**[§], **Anne C. Moorman, MPH, RN**[§], **Philip R. Spradling, PhD**[§], **Eyasu Teshale, PhD**[§], **Joseph A. Boscarino, MD**[‖], **Yihe G. Daida, PhD**[¶], **Mark A. Schmidt, MD**[#], and **Mei Lu, PhD**[†] **For the CHeCS investigators**

[*]Division of Gastroenterology and Hepatology, Henry Ford Health System, Detroit, MI

[†]Department of Public Health Sciences, Henry Ford Health System, Detroit, MI

[‡]Center for Health Policy and Health Services Research, Henry Ford Health System, Detroit, MI

[§]Division of Viral Hepatitis, Centers for Disease Control and Prevention, Atlanta, GA

[‖]Center for Health Research, Geisinger Health System, Danville, PA

[¶]Center for Health Research, Kaiser-Permanante Hawaii, Honolulu, HI

[#]Center For Health Research, Kaiser-Permanante Northwest, Portland, OR

Abstract

Goals—To determine the impact of geography and patient characteristics on hepatitis C virus (HCV) genotype and subtype distribution in a large sample of patients under routine clinical care

Background—HCV genotype impacts disease course and response to treatment. Although several studies have reported genotype distribution within specific US populations, there are no comprehensive descriptions in large, geographically diverse cohorts.

Study—Using data from the Chronic Hepatitis Cohort Study, we present the distribution of HCV genotypes (GT) and subtypes (ST) among a racially diverse cohort of over 8000 HCV-infected patients from four large US health systems.

Results—Genotype distribution varied significantly by geographic and demographic factors. In age-adjusted analyses, African American patients had significantly higher prevalence of GT1 (85%) than other racial categories, largely driven by a markedly higher proportion of GT1 subtype b (~34%) than in Asian/other (24%) and white (21%) patients. GT3 represented an increasing proportion of infections as birth decade progressed, from 4% in patients born before 1946 to 18% of those born after 1976. Within the cohort of “living/uncured” patients, highly elevated alanine aminotransferase (> 2 times the upper limit of normal) was significantly more common in GT3

Address correspondence to: Stuart C. Gordon, MD, Henry Ford Hospital, 2799 West Grand Blvd, Detroit, MI 48202, sgordon3@hfhs.org.

The remaining authors declare that they have nothing to disclose.

Supplemental Digital Content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal’s Website, www.jcge.com.

patients, whereas Fibrosis-4 Index scores indicative of cirrhosis were most common in the combined group of GT4&6 patients.

Conclusion—Distribution of HCV genotypes and subtypes in the United States is more variable than suggested by previous national-level estimates and single-center studies. “Real-world” prevalence data may improve targeting of prevention, screening, and treatment efforts for hepatitis C.

Keywords

HCV; genotype 1b; genotype 3; racial disparities; ALT

The hepatitis C virus (HCV) demonstrates remarkable genetic diversity, which complicates the development of effective treatments and potential vaccines. Even in the era of direct-acting all-oral antiviral (DAA) therapies, treatment selection and effectiveness depend on viral genotype (GT) and subtype (ST). However, there have been few studies of HCV GT/ST distribution in the United States, and these have been confined to single health systems or limited by small sample sizes, large proportions of male patients, and incomplete description of subtypes.[1–3] Furthermore, these studies also have not examined how various clinical factors, such as liver biochemistries, may vary across genotype and subtype. More comprehensive genotype and subtype data are necessary to guide HCV prevention, screening, and treatment efforts in the United States.

The Chronic Hepatitis Cohort Study (CHeCS) is a longitudinal study of hepatitis patients from four large US health systems, comprising a geographically and racially diverse cohort of over 10,000 HCV-infected patients. CHeCS has over 10 years’ of extensively annotated patient data, which permits us to report the real-world distribution of HCV GT/ST across site and patient characteristics, as well as to describe how clinical factors vary by GT/ST. In addition, we provide the genotype/subtype distribution of the living/uncured patients in our cohort—that is, patients who remain eligible for treatment with emerging DAA therapies.

METHODS

Study Population

CHeCS is an observational multicenter study that includes adult patients (18 y and above) from four large health systems. The study follows all guidelines of the US Department of Health and Human Services regarding the protection of human subjects; protocols are reviewed annually by the institutional review board at each study site—Geisinger Health System (GHS), Danville, PA; Henry Ford Health System (HFHS), Detroit, MI; Kaiser-Permanente Hawai’i (KPHI), Honolulu, HI; and Kaiser-Permanente Northwest (KPNW), Portland, OR. Written informed consent was waived due to the de-identified nature of the data. CHeCS study methods have been previously described.[4]

Patient demographic information was collected at the time of entry into the cohort. Clinical data (including laboratory results for imputation of the Fibrosis-4 [FIB4] score[5]) were collected for the most recent encounter before December 2014. FIB4, a serum marker of hepatic fibrosis, was classified into one of three categories,[6] using previously validated

cutoffs of ≤1.21 (for Metavir stages F0-2; “no to moderate fibrosis”), 1.21≥5.88 (F3-4; “advanced fibrosis”), and >5.88 (F4; cirrhosis). Charlson-Deyo comorbidity indices[7] were imputed from electronic health record data for one year before each patient’s last encounter through December 31, 2014 to estimate their chronic disease burden. Detailed antiviral medication data were collected via chart abstraction. Treated patients were classified as having achieved sustained virological response (SVR) if laboratory results ≥12 weeks post-therapy showed undetectable viral loads; those who did not achieve SVR were classified as “treatment failure.” Death data were collected through various resources including the National Death Index and Social Security Death Index.

HCV genotype was collected from electronic capture of laboratory results, supplemented with chart review. Patients with more than one or inconsistent genotypes in the medical record were classified by the genotype reported first. For GT1 and 2, subtypes were classified as “a,” “b,” or “other” ([o], >1 subtype or not reported in the medical record). Because of low frequency, subtypes for other genotypes were not reported. The sole patient with GT5 was excluded from analyses.

Statistical Analyses

Descriptive summary statistics are presented as counts and percentages (row percentages for Tables 1A and 2A, column percentages for Tables 1B and 2B). Chi-squared tests were used to test the population difference between genotyped and nongenotyped patients, and among GT/ST categories. Bar charts were used to illustrate proportional distribution of GT/ST by patient demographic factors and proportional frequency of clinical factors across GT/ST.

We used generalized logit models to determine whether GT/ST distributions differed by geography (clinical site) and patient characteristics (sex, race, and birth decade). Multivariable analyses were performed to adjust for possible confounding.[4] Pairwise GT/ST comparisons were performed if there was overall significance at the criteria of 0.05. Adjusted probabilities (percent) of each GT/ST were estimated from these models using the marginal standardization method.[8,9] Briefly, this method calculates the estimated probabilities separately for each level of variables in the model, and then combined as a weighted average for each level of variable of interest. 95% confidence intervals (CI) were estimated using the Bootstrap approach.[10] Differences between two groups are significant if the CIs do not overlap. A similar analytic approach was used to study the distribution of clinical factors across GT/ST, adjusting for patient characteristics (such as race and birth year).

RESULTS

HCV genotype was available for 8140 (75%) of the 10,738 HCV-infected patients in our cohort, after exclusion of the sole patient with genotype 5. There were a number of significant differences between patients with and without genotype data (Supplemental Table 1, Supplemental Digital Content 1, http://links.lww.com/JCG/A351). As genotyping is often performed before initiation of antiviral therapy, patients with HCV genotyping were more likely to receive HCV treatment than those without genotyping (46% vs. 23%). Patients with available data were also significantly more likely to have private insurance and

to have fewer comorbid conditions. Notably, markers of the severity of liver disease [alanine aminotransferase (ALT) levels and FIB4 indices] did not differ significantly between patients with and without genotype data. Patients co-infected with either human immunodeficiency virus or hepatitis B virus represented only 1% to 3% of all genotyped patients. Because rates did not differ significantly by GT/ST, these patients were included in all analyses.

Univariate Analyses: All Genotyped Patients

Table 1 and Figure 1 present the distribution of GT/ST across patient characteristics for all genotyped patients. GT1 was most common (75.4%), followed by 2 (12.6%) and 3 (10.2%); GT4 (1.5%) and 6 (0.3%) were much less prevalent.

However, chi-squared tests showed significant differences in the distribution of GT/ST across all demographic variables, as well as by site ($P < 0.01$, except sex; $P = 0.022$). For example, the proportion of patients with GT1 ranged from 65% (KPNW, Portland, OR) to 84% (HFHS, Detroit, MI). Black/African American patients had significantly higher prevalence of GT1 (over 93%) than other racial categories; this was largely driven by a markedly larger proportion of GT1b than in Asian/other and white patients. Likewise, as birth year increased (age decreased), the proportion of GT1b also decreased, from 41% in patients born before 1946 to <10% in patients born after 1975. This trend was reflected by changes in the proportion of GT3, which became more common as age decreased, representing only 3% of infections in the oldest patients but 18% of infections among the youngest patients.

Multivariate Analyses and Adjusted Probability: All Genotyped Patients

Birth decade, patient race, and study site were all independently associated with genotype distribution ($P < 0.01$). Although genotype distribution varied significantly by sex in the univariate analysis, the significant association did not persist in the multivariate analysis. Figures 2–4 display the adjusted probabilities for each GT/ST across CHeCS sites, patient race, and birth decade, respectively. 95% CIs are included to permit comparisons of significant differences within categories and between GT/ST.

Sites/Geography

After adjustment for differences in the racial and age makeup of each CHeCS health system, the GT/ST distribution remained significantly different across sites (Fig. 2). As in the univariate analysis, the largest proportion of infections were GT1a (41.1% to 48.6%). However, HFHS (Detroit, MI) and KPHI (Honolulu, HI) had significantly larger proportions (4.8% to 7.5% higher) of GT1a than GHS and KPNW (Dansville, PA and Portland OR, respectively). GT3 demonstrated a similar geographic pattern (higher proportions at HFHS and KPHI; lower proportions at GHS and KPNW).

HFHS demonstrated a significantly higher adjusted probability of GT1b (27.6% vs. 20.1% to 22.1%) and GT4 (2.5% vs. 0.4% to 1.5%) compared with other sites; adjusted probabilities of 2b and 2o were significantly lower (5.0% vs. 8.3% to 11.5%, and 1.0% vs. 3.2% to 6.4%, respectively). KPNW demonstrated a roughly 1.5- to 2-fold higher

probability of GT2b (11.5%) and 2o (6.4%) than other sites, and a 4- to 8-fold higher probability of GT6 (0.8% vs. 0.1% to 0.2%).

Race

After adjustment for differences in the age distribution of patients in each racial category, GT/ST distribution remained significantly different across patient race (Fig. 3). Even after age-adjustment, black/African American patients demonstrated 8% to 12% higher adjusted probability of GT1a (51.3% vs. 39.2% to 43.3%) and 10% to 13% higher probability of GT1b (33.6% vs. 20.5% to 23.6%) than Asian/other or white patients; this was reflected in significantly lower probabilities of GT2b, 2o, 3, and 4 than other racial groups.

GT6 was confined exclusively to Asian/other patients (3.1% vs. 0%). Asian/other patients also demonstrated the highest adjusted probability of GT4 (4.9% vs. 0.4% to 1.6%), and less overall variability in GT/ST distribution than other racial groups. White patients demonstrated the highest adjusted probabilities of GT2b (9.8%) and 3 (13.1%).

Birth Decade

After adjustment for differences in the racial and geographic distribution of patients in our sample, the distribution of GT/ST remained significantly different across birth decades (Fig. 4). As in the unadjusted analyses, GT1a represented the largest proportion of infections across all birth decades, except for patients born before 1946, who were equally likely to have GT1b. The adjusted probability of 1b infections decreased steadily and significantly across birth decades, from 35% in patients born before 1946 to <11% in patients born after 1975. GT2b and 2o demonstrated a similar but nonsignificant pattern, although these differences were significant when comparing the oldest (< 1946) and youngest (≥1976) patients.

The adjusted probabilities of GT1a, 3, and 4 increased across each birth decade, from oldest to youngest patients. Proportion of GT1a rose from 32.7% to 55.3%, a 1.7-fold increase, while the proportion of GT4 rose from 1.4% to 4%, a roughly 3-fold increase. GT3 demonstrated the largest proportional increase, from 3.6% in patients born before 1946 to 17.8% in patients born after 1976, a 5-fold increase. Among patients born after 1976, GT1a and 3 were the most common, representing almost three-quarters of all infections among the youngest patients.

Notably, we observed significant differences between the two decades comprised by the “Baby Boomer” birth cohort, a group that is often analyzed as a single unit. Older patients in this cohort (born 1946 to 1955) were significantly more likely than their younger counterparts (born 1956 to 1965) to have GT1b (adjusted probability of 27% vs. 19.5%). Younger cohort members were significantly more likely to have GT1a (46.8% vs. 43.2%) and GT3 (12.9% vs. 7.7%) than older cohort members.

Univariate Analysis: Living/Uncured Patients

Unadjusted treatment receipt and vital status across GT/ST are described in Table 1 and Figure 1. Because of the necessity of using complex time-to-event analyses to determine the

impact of age and other confounders upon these outcomes, multivariate analyses were not performed. The cohort of 5138 patients who were still living and had not achieved SVR as of December 2014 is described in Table 2. In general, distribution of genotypes across "living/uncured" patients was similar to that of the "all genotyped" patients, with the exception of a slight increase (2% to 7%) in the proportion of GT1a patients across all demographic characteristics. In unadjusted analyses, clinical factors varied across GT/ST (Fig. 5); ALT, FIB4 were selected for further analyses.

Multivariate Analysis: Living/Uncured Patients

Although the relationship between GT/ST and comorbidity indices did not remain significant after adjustment (Fig. 5), the adjusted probability of ALT and FIB4 varied significantly by genotype (Fig. 6). Among living/uncured patients with available ALT data (n = 3715), roughly half (49.0% to 63.6%) had "normal" ALT levels. GT3 patients demonstrated the highest probability of highly elevated upper limit of normal (≥2*ULN) ALT (21.1% vs. 11.6% to 14.1%). This difference was significant for GT1a and 1b, and approached significance for GT1o and 2; GT4&6 displays a wide CI due to small sample size (n = 12). GT3 patients were also the least likely to have "normal" ALT (49.0% vs. 53.3% to 63.6%, significantly lower than 1a, 1b, and 2). GT2 patients were more likely to have normal ALT and less likely to have "elevated" ALT (normal <2*ULN) compared with other patients with other GT/STs; this difference was significant for GT1a and 3.

We have previously shown that a FIB4 cutoff of >5.88 reliably predicts cirrhosis in this cohort.[6] Among living/uncured patients with available FIB4 data (n = 3173), GT4&6 patients demonstrated the largest proportion of patients with FIB4-defined cirrhosis; even with a wide CI, this proportion was significantly larger than that of other GT/STs except GT3 and 1b. GT3 patients also demonstrated a high probability of FIB4-indicated cirrhosis (19%); this was significantly higher than GT1a and 2. The majority of patients of all genotypes (~55% to 67%, except GT4&6) had FIB4 values of 1.21 to 5.88, suggesting that moderate-to-severe fibrosis is common among the cohort of patients who remain candidates for DAA therapy.

DISCUSSION

In this cohort of over 10,000 US patients, we observed marked variation in genotype and subtype distribution across site-level and patient-level characteristics. Results from previous studies confined to single-state populations (California[1] and Texas[3]) have suggested that genotype distribution varies by both geography and demographics, although it is not possible to directly compare these cohorts. These studies are also limited by a lack of clinical data or restricted to univariate analyses. We applied extensive medical record data from four large health systems to both univariate and multivariate analyses to generate a comprehensive picture of both genotype/subtype distribution and associated clinical characteristics.

Although "Baby Boomer" HCV-infected patients are often analyzed collectively, our data suggest that this is not a homogenous group. In both unadjusted and adjusted analyses, patients born in the second decade of this cohort (1956 to 1965) had significantly lower proportions of GT1b and higher proportions of GT1a and 3 than their older counterparts

(born from 1946 to 1955); in general, younger cohort members resembled patients born the following decade (from 1966 to 1975, no significant differences in proportions between these two birth decades) more than they resembled those born from 1946 to 1955.

Consistent with these findings, as birth decades progressed, the proportions of GT1b significantly decreased while the proportions of GT1a and 3 increased; among patients born 1976 and after, the proportion of GT3 infections was five times higher than that among patients born before 1946. This finding is consistent with a large, laboratory-based study that found GT3 patients to be significantly younger than those with other genotypes.[11] This genotype shift likely represents changes in patterns of HCV exposure over time in the United States—from medical exposure (associated with GT1b) to injection drug use (frequently associated with GT1a and 3).[12]

This increasing proportion of GT3 infection among younger patients has important epidemiological and clinical considerations.[13] In analyses adjusted for differences in race and age, we found significant differences in ALT values between genotypes. Among our living/uncured patient cohort, GT3 patients had the highest rates of “highly elevated” ALT; over 21% of GT3 patients had ALT values >2 times the upper limit of normal. We are not aware that this observation has been previously reported in a primarily HCV mono-infected cohort. GT3 is associated with aggressive progression of fibrosis and cirrhosis,[14] likely contributing to the increased risk of hepatocellular carcinoma (HCC) that we and others have reported in GT3 patients.[15,16] Although high rates of hepatic steatosis and diabetes have been implicated in this rapid progression of liver disease, a recent in vitro study found that GT3 also provokes a uniquely inflammatory hepatic immune response; the elevated ALT values we observed in GT3 patients may reflect increased necroinflammatory activity.[17] Additional research exploring the implications of this observation is warranted.

Likewise, there is a significant burden of fibrosis among our entire cohort of living/uncured patients; roughly 70% to 80% of living/uncured patients have FIB4 indices >1.21 (at least moderate fibrosis). Although direct comparisons are not possible (all FIB4 patients have ALT values but not viceversa), the majority of living/uncured patients demonstrated “normal” ALT values, but only 20% to 30% of patients demonstrated FIB4 values consistent with “no to mild” fibrosis (< 1.21).[6] In fact, over 25% of GT4&6 patients and 19% of GT3 patients demonstrated FIB4-indicated cirrhosis—despite the relatively younger age profile of the latter. Clinical awareness of the variability of indicators of the severity of liver inflammation and fibrosis as a function of GT/ST, even after controlling for age and race, may influence both screening and treatment decisions.

GT/ST distribution varied significantly by race in both univariate analyses as well as multivariate analyses that adjusted for differences in age distribution across the 3 racial categories used. African American patients overwhelmingly demonstrated infection with GT1—most notably an adjusted probability of 1b (33%) roughly 1.5 times that of white (21%) and Asian (24%) patients. We note that these rates were consistent between the full cohort and the living/uncured cohort. A similarly high rate of 1b in black/African American patients was shown in a recent study from a single health center in Texas,[3] a region not represented in our sample. To our knowledge, an epidemiological explanation for the

dramatic over-representation of this subtype within African Americans has not been reported. However, given that African Americans have been less likely to receive interferon-based HCV treatment—for reasons that include higher rates of comorbidities/contraindications and lower response rates compared with other racial/ethnic groups[18–21]—the efficacy of newer regimens in treating this subtype[22] may address some of the disparities in treatment and long-term outcomes observed in these patients. We also observed that GT6 infection was confined almost entirely to Asian patients. Given the relatively higher prevalence of GT6 in China and other parts of Asia,[23] it is likely that the GT6 patients in our cohort were infected before arrival in the United States.

Recent worldwide HCV genotype distribution estimates considered the US or North America as a single unit.[23,24] However, our data show that genotype distribution varies considerably across study sites, even after adjusting for differences in the age and racial makeup of each cohort. Although previous US-based studies have demonstrated regional differences in genotype distribution, variation across regions was lower (< 10%),[11] and did not control for differences in age and racial makeup of each region. Other studies of genotype distribution in the US were limited by small sample sizes[2,25] or were confined to a single health system[1] or geographic region.[25] Although our results are generally consistent with these reports, our analysis extends these findings by including a broader spectrum of genotypes/subtypes and their association with clinical characteristics; we also performed multivariate analyses that allowed us to adjust these analyses for race, birth decade, and clinical site. A considerable strength of this analysis is that CHeCS comprises sites spanning the continental US as well as Hawai'i. Notably, relative proximity did not appear to influence genotype distribution; in general, adjusted probabilities for each GT/ST were most similar between HFHS (Detroit, MI) and KPHI (Honolulu, HI). Although this study is not designed to investigate factors related to the local epidemiology of HCV, our data demonstrate that GT distribution in a single community or health system may not be generalizable to the US as a whole.

We acknowledge that there is some question regarding the future clinical relevance of HCV genotype as pangenotypic treatments are anticipated to become the standard of care. Nevertheless in 2017, genotype—coupled with previous treatment experience and degree of fibrosis—remains an important parameter when determining treatment. We believe that a comprehensive illustration of pre-DAA era HCV genotype distribution in the US, particularly among the pool of HCV patients who remain candidates for DAA treatments, remains important to the clinician. As the more easily treated genotypes become less common, the proportion of "difficult to treat" genotypes will increase; as a result, genotype-specific regimens will likely remain part of the clinician's repertoire.

Likewise, treatment remains but one clinical indication for genotype identification. For example, genotype may be used to determine whether a "cured" patient has been re-infected or is experiencing viral reemergence. As we and others have shown, clinical outcome and hepatocellular cancer risk varies by genotype; for example, we have previously shown that genotype 3 is associated with increased risk of HCC compared with other genotypes.[26] Our data show demographic shifts in genotype frequency (from 1b among older patients to 3 among younger patients). The increasing proportion of genotype 3 among younger

individuals—presumably with shorter disease duration and subsequently less fibrosis—should heighten public awareness and reinforce efforts toward earlier treatment. Left untreated, these patients are likely to progress rapidly to advanced fibrosis and an increased risk for HCC. Differences in long-term outcomes, even among patients who achieve SVR, may also vary by genotype.

Another limitation of our study is a lack of data regarding source of infection. A large subset of CHeCS patients responded to a survey that included questions about how they acquired HCV; however, the majority of patients did not know or declined to answer this question (data not shown). This lack of data regarding the source of infection limits our ability to investigate underlying reasons for the observed differences in GT/ST distribution as well as our ability to estimate duration of infection. The cross-sectional analysis also limits our ability to draw causal inferences, and some small cell counts, particularly within our living/uncured cohort, reduce statistical power to observe some significant differences.

The timeline of our data collection—through December 2014—means that we have little data regarding patients who have been treated since the widespread implementation of DAA treatment regimens. However, roughly one-quarter to one-third of all living patients in this cohort have previously failed treatment, further increasing their risk for cirrhosis and HCC.[15] Although proportions varied by genotype, our entire cohort of living/uncured patients demonstrate high levels of moderate-to-severe fibrosis and cirrhosis, as indicated by FIB4 scores. Although a number of DAA regimens are now available, access to these therapies has been limited, and there remains a large pool of patients who may be considered priority candidates for therapy.[27]

In conclusion, the overall distribution of HCV genotypes and subtypes in the United States reported in this analysis largely reflects previous estimates, but these estimates paint an incomplete picture. On the basis of comprehensive data from over 8000 HCV-infected patients from four large health systems, we show that genotype distribution varies significantly by geography and demographics, with genotype-distinct clinical features, creating unique subpopulations. The current detailed characterization demonstrates wide variation in the proportion of HCV genotypes within a large real-world population and may inform the development and optimization of US HCV eradication efforts.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The CHeCS Investigators include the following investigators and sites: Scott D. Holmberg, Eyasu H. Teshale, Philip R. Spradling, Anne C. Moorman, Jim Xing, and Yuna Zhong, Division of Viral Hepatitis, National Centers for HIV, Viral Hepatitis, STD, and TB Prevention (NCHHSTP), Centers for Disease Control and Prevention (CDC), Atlanta, GA; Stuart C. Gordon, David R. Nerenz, Mei Lu, Lois Lamerato, Jia Li, Loralee B. Rupp, Nonna Akkerman, Talan Zhang, Sheri Trudeau, Yueren Zhou, and Kuan-Han Wu. Henry Ford Health System, Detroit, MI; Joseph A. Boscarino, Zahra S. Daar, and Robert E. Smith, Department of Epidemiology and Health Services Research, Geisinger Health System, Danville, PA; Yihe G. Daida, Connie Mah Trinacty, Jonathan W. Lai, and Carmen P. Wong, The Center for Health Research, Kaiser-Permanente Hawaii, Honolulu, HI; Mark A. Schmidt Judy L. Donald, The Center for Health Research, Kaiser-Permanente Northwest, Portland, OR.

Henry Ford Health System receives funding for CHeCS from the Centers for Disease Control and Prevention and from Gilead Sciences. CHeCS was previously funded through May 2016 by the CDC Foundation, which received grants from AbbVie; Genentech, A Member of the Roche Group; Gilead Sciences; Janssen Pharmaceuticals Inc. and Vertex Pharmaceuticals; past partial funders include Bristol-Myers Squibb. Granting corporations do not have access to CHeCS data and do not contribute to data analysis or writing of manuscripts.

S.C.G.: receives grant/research support from AbbVie Pharmaceuticals, Bristol-Myers Squibb, Conaturs, CumaBay, Exaclenz BioScience, Gilead Pharmaceuticals, Intercept Pharmaceuticals, and Merck. He is also a consultant/advisor for AbbVie Bristol-Myers Squibb, CVS Caremark, Gilead Pharmaceuticals, and Merck.

References

1. Manos MM, Shvachko VA, Murphy RC, et al. Distribution of hepatitis C virus genotypes in a diverse US integrated health care population. J Med Virol. 2012; 84:1744–1750. [PubMed: 22997077]
2. Nainan OV, Alter MJ, Kruszon-Moran D, et al. Hepatitis C virus genotypes and viral concentrations in participants of a general population survey in the United States. Gastroenterology. 2006; 131:478–484. [PubMed: 16890602]
3. Xie Y, Garza G, Dong J. Hepatitis C virus genotype and subtype distribution in patient specimens tested at the University of Texas Medical Branch, Galveston, Between January 2011 and November 2014. Lab Med. 2016; 47:112–118. [PubMed: 26995188]
4. Moorman AC, Gordon SC, Rupp LB, et al. Baseline characteristics and mortality among people in care for chronic viral hepatitis: the chronic hepatitis cohort study. Clin Infect Dis. 2013; 56:40–50. [PubMed: 22990852]
5. Teshale E, Lu M, Rupp LB, et al. APRI and FIB-4 are good predictors of the stage of liver fibrosis in chronic hepatitis B: the Chronic Hepatitis Cohort Study (CHeCS). J Viral Hepat. 2014; 21:917–920. [PubMed: 25131445]
6. Li J, Gordon SC, Rupp LB, et al. The validity of serum markers for fibrosis staging in chronic hepatitis B and C. J Viral Hepat. 2014; 21:930–937. [PubMed: 24472062]
7. Deyo RA, Cherkin DC, Ciol MA. Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases. J Clin Epidemiol. 1992; 45:613–619. [PubMed: 1607900]
8. Joffe MM, Greenland S. Standardized estimates from categorical regression models. Stat Med. 1995; 14:2131–2141. [PubMed: 8552892]
9. Muller CJ, MacLehose RF. Estimating predicted probabilities from logistic regression: different methods correspond to different target populations. Int J Epidemiol. 2014; 43:962–970. [PubMed: 24603316]
10. Efron B. The Jackknife, the Bootstrap, and Other Resampling Plans. Stanford, CA: Stanford University; 1982.
11. Germer JJ, Mandrekar JN, Bendel JL, et al. Hepatitis C virus genotypes in clinical specimens tested at a national reference testing laboratory in the United States. J Clin Microbiol. 2011; 49:3040–3043. [PubMed: 21613437]
12. Jacka B, Applegate T, Krajden M, et al. Phylogenetic clustering of hepatitis C virus among people who inject drugs in Vancouver, Canada. Hepatology. 2014; 60:1571–1580. [PubMed: 25042607]
13. Feng S, Lai JC. Expanded criteria donors. Clin Liver Dis. 2014; 18:633–649. [PubMed: 25017080]
14. Probst A, Dang T, Bochud M, et al. Role of hepatitis C virus genotype 3 in liver fibrosis progression—a systematic review and meta-analysis. J Viral Hepat. 2011; 18:745–759. [PubMed: 21992794]
15. Lu M, Li J, Rupp LB, et al. Hepatitis C treatment failure is associated with increased risk of hepatocellular carcinoma. J Viral Hepat. 2016; 23:718–729. [PubMed: 27028626]
16. Gondeau C, Pageaux GP, Larrey D. Hepatitis C virus infection: are there still specific problems with genotype 3? World J Gastroenterol. 2015; 21:12101–12113. [PubMed: 26576095]
17. Mitchell AM, Stone AE, Cheng L, et al. Transmitted/founder hepatitis C viruses induce cell-type- and genotype-specific differences in innate signaling within the liver. MBio. 2015; 6:e0251032510–14.

18. Khokhar OS, Lewis JH. Reasons why patients infected with chronic hepatitis C virus choose to defer treatment: do they alter their decision with time? Dig Dis Sci. 2007; 52:1168–1176. [PubMed: 17357838]
19. Borum ML, Igiehon E, Shafa S. African Americans may differ in their reasons for declining hepatitis C therapy compared to non-African Americans. Dig Dis Sci. 2009; 54:1604. Author reply 5. [PubMed: 19399619]
20. Melia MT, Muir AJ, McCone J, et al. Racial differences in hepatitis C treatment eligibility. Hepatology. 2011; 54:70–78. [PubMed: 21488082]
21. Hare CB, Morris JA, Chu A, et al. Comparison of characteristics of treated and non-treated patients with hepatitis C infection. Pharmacoepidemiol Drug Saf. 2006; 15:71–76. [PubMed: 16136612]
22. Manns M, Pol S, Jacobson IM, et al. All-oral daclatasvir plus asunaprevir for hepatitis C virus genotype 1b: a multinational, phase 3, multicohort study. Lancet. 2014; 384:1597–1605. [PubMed: 25078304]
23. Messina JP, Humphreys I, Flaxman A, et al. Global distribution and prevalence of hepatitis C virus genotypes. Hepatology. 2015; 61:77–87. [PubMed: 25069599]
24. Gower E, Estes C, Blach S, et al. Global epidemiology and genotype distribution of the hepatitis C virus infection. J Hepatol. 2014; 61(suppl):S45–S57. [PubMed: 25086286]
25. Dias PT, Hahn JA, Delwart E, et al. Temporal changes in HCV genotype distribution in three different high risk populations in San Francisco, California. BMC Infect Dis. 2011; 11:208. [PubMed: 21810243]
26. Lu M, Li J, Zhang T, et al. Serum biomarkers indicate long-term reduction in liver fibrosis in patients with sustained virological response to treatment for HCV infection. Clin Gastroenterol Hepatol. 2016; 14:1044.e3. [PubMed: 26804385]
27. Xu F, Leidner AJ, Tong X, et al. Estimating the number of patients infected with chronic HCV in the United States who meet highest or high-priority treatment criteria. Am J Public Health. 2015; 105:1285–1289. [PubMed: 25973816]

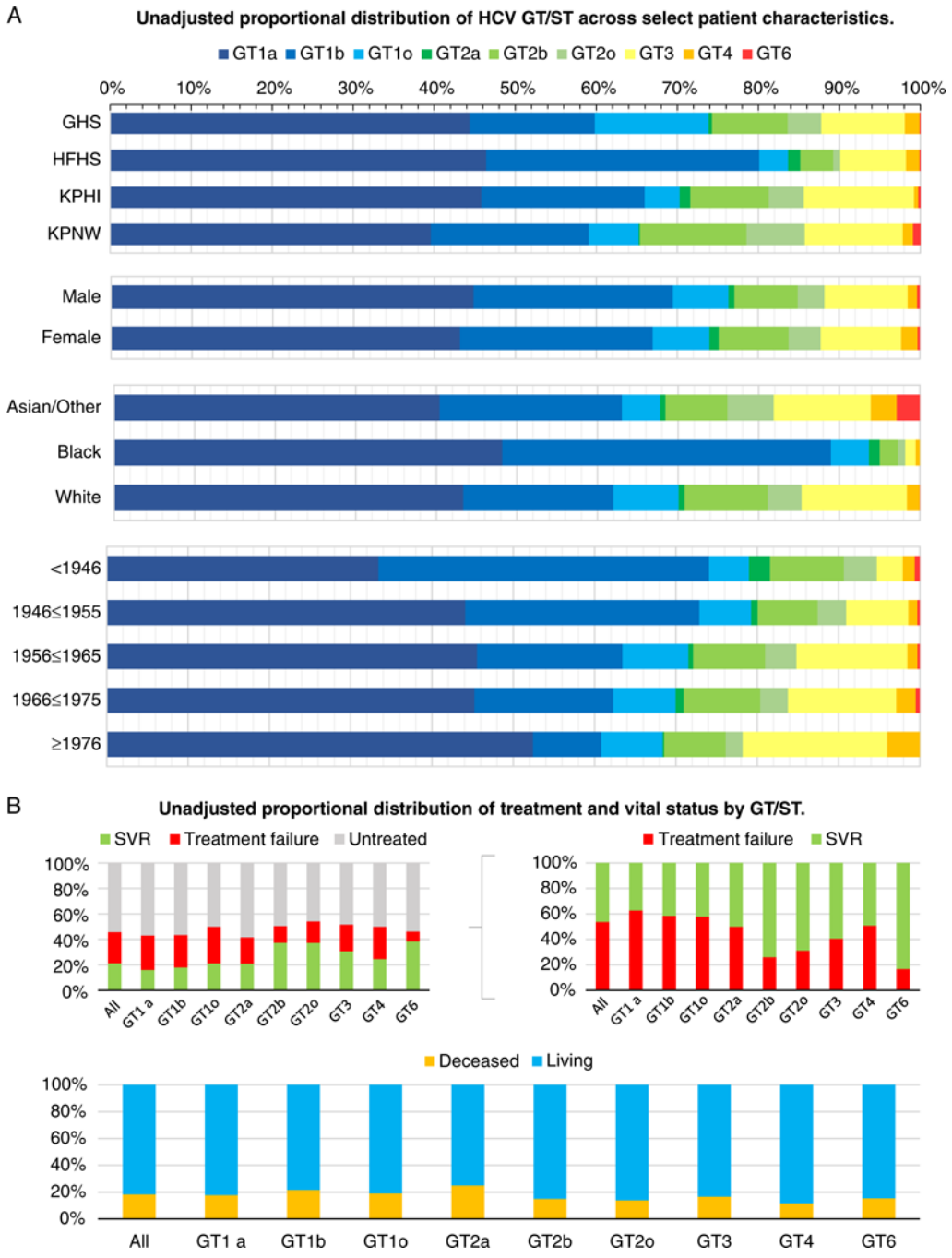


FIGURE 1.

All genotyped patients (N = 8140). A, Unadjusted proportional distribution of HCV GT/ST by site and patient-level characteristics. B, Unadjusted proportional distribution of treatment and vital status by GT/ST. GHS indicates Geisinger Health System; GT, genotype; HCV, hepatitis C virus; HFHS, Henry Ford Health System; KPHI, Kaiser-Permanente Hawai'I; KPNW, Kaiser-Permanente Northwest; SVR, sustained virological response; ST, subtype.

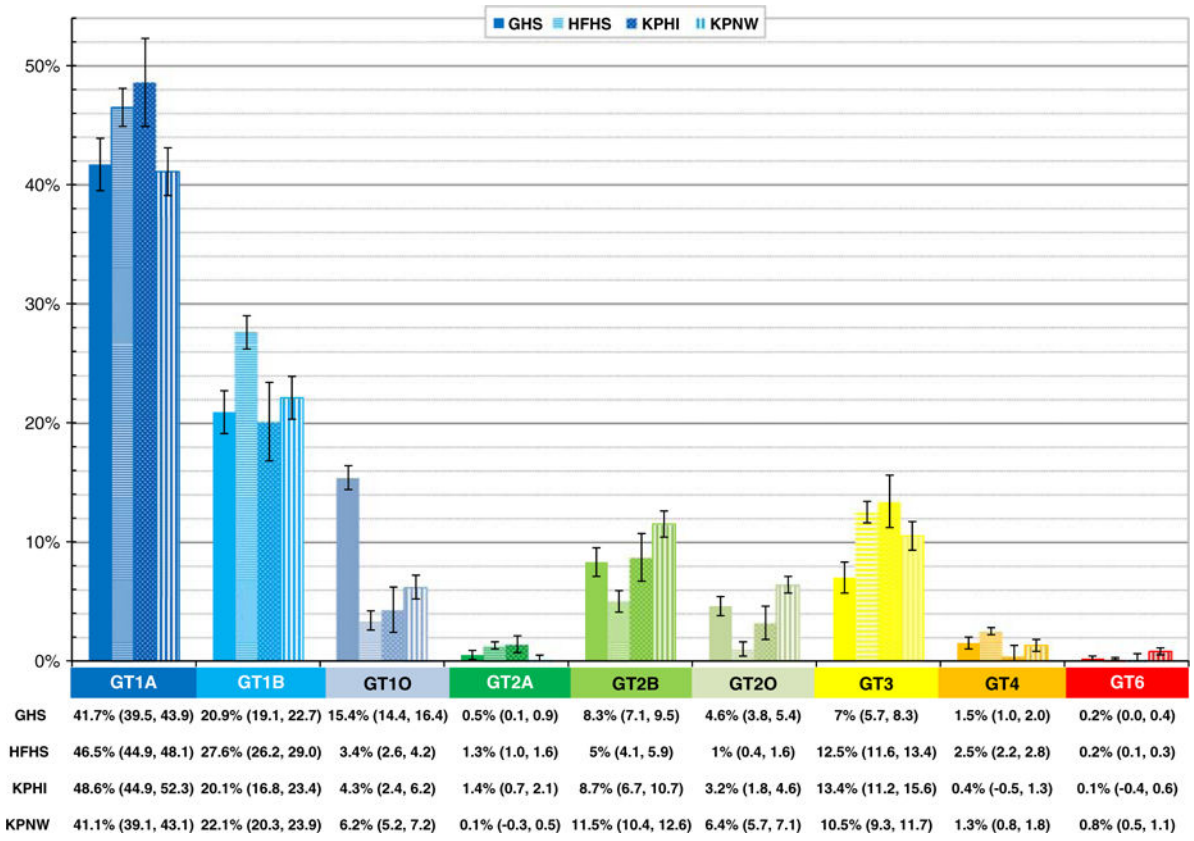


FIGURE 2.

All genotyped patients (N = 8140), adjusted probability of GT/ST by CHeCS site. Category (site) percentages sum to 100. GHS indicates Geisinger Health System; GT, genotype; HFHS, Henry Ford Health System; KPHI, Kaiser-Permanente Hawai'I; KPNW, Kaiser-Permanente Northwest; ST, subtype.

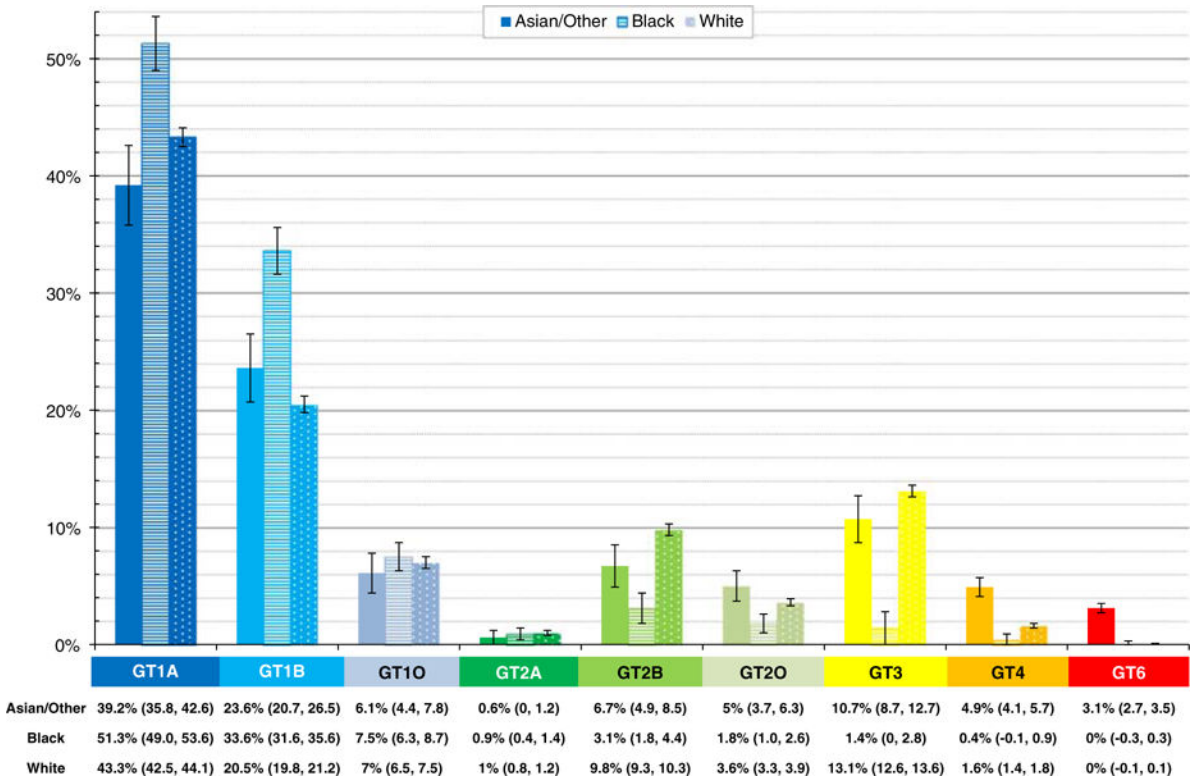
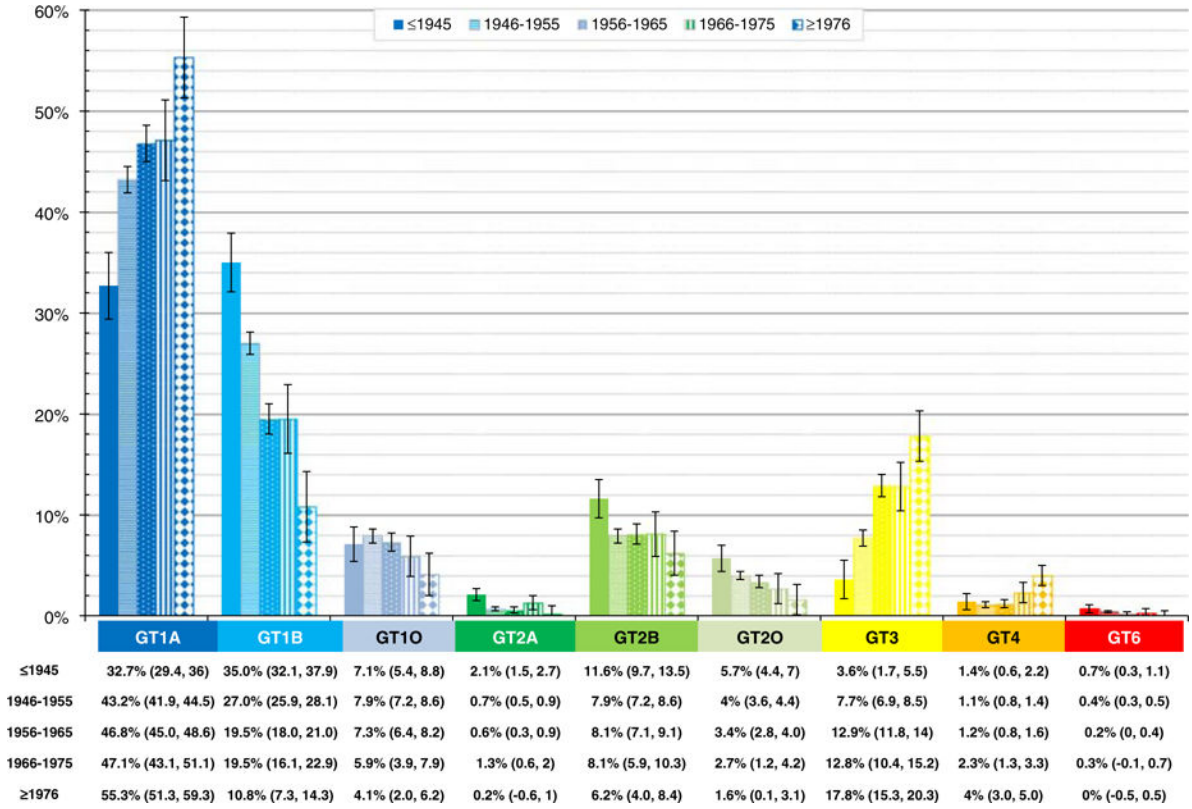


FIGURE 3.

All genotyped patients (N = 8140), adjusted probability of GT/ST by race. Category (race) percentages sum to 100. GT indicates genotype; ST, subtype.



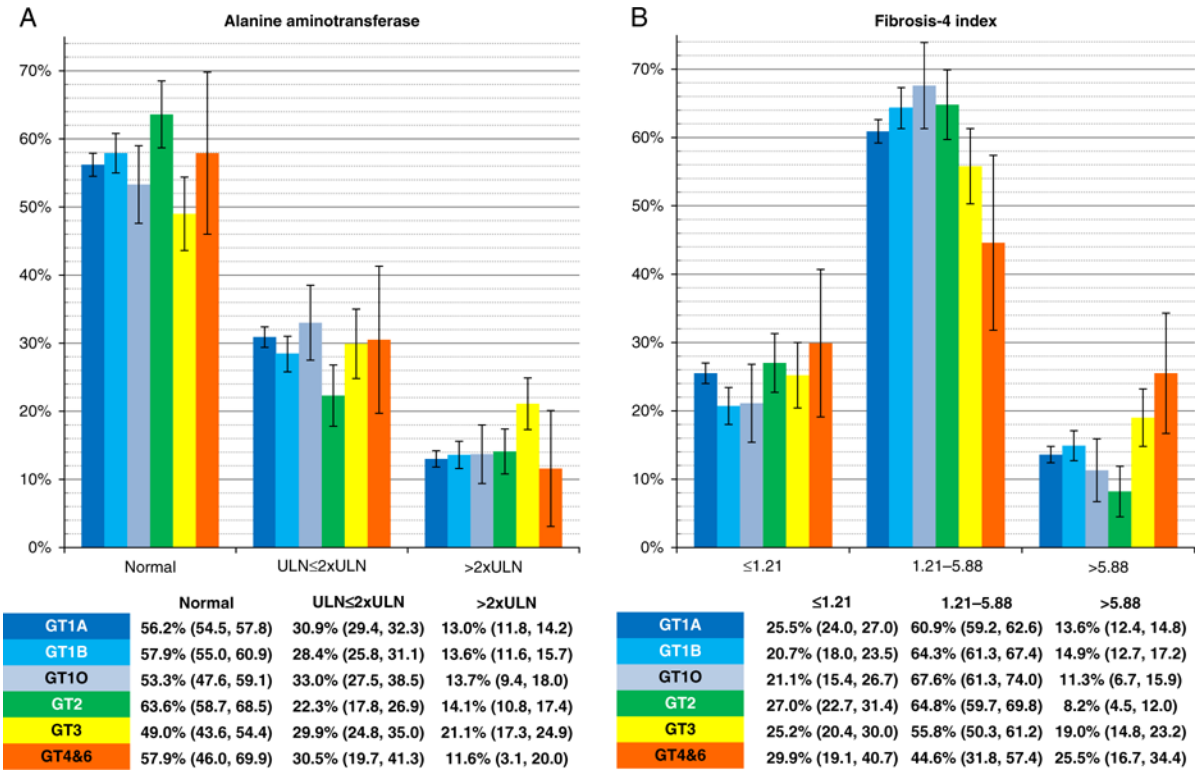
	GT1A	GT1B	GT1O	GT2A	GT2B	GT2O	GT3	GT4	GT6
≤1945	32.7% (29.4, 36)	35.0% (32.1, 37.9)	7.1% (5.4, 8.8)	2.1% (1.5, 2.7)	11.6% (9.7, 13.5)	5.7% (4.4, 7)	3.6% (1.7, 5.5)	1.4% (0.6, 2.2)	0.7% (0.3, 1.1)
1946-1955	43.2% (41.9, 44.5)	27.0% (25.9, 28.1)	7.9% (7.2, 8.6)	0.7% (0.5, 0.9)	7.9% (7.2, 8.6)	4% (3.6, 4.4)	7.7% (6.9, 8.5)	1.1% (0.8, 1.4)	0.4% (0.3, 0.5)
1956-1965	46.8% (45.0, 48.6)	19.5% (18.0, 21.0)	7.3% (6.4, 8.2)	0.6% (0.3, 0.9)	8.1% (7.1, 9.1)	3.4% (2.8, 4.0)	12.9% (11.8, 14)	1.2% (0.8, 1.6)	0.2% (0, 0.4)
1966-1975	47.1% (43.1, 51.1)	19.5% (16.1, 22.9)	5.9% (3.9, 7.9)	1.3% (0.6, 2)	8.1% (5.9, 10.3)	2.7% (1.2, 4.2)	12.8% (10.4, 15.2)	2.3% (1.3, 3.3)	0.3% (-0.1, 0.7)
≥1976	55.3% (51.3, 59.3)	10.8% (7.3, 14.3)	4.1% (2.0, 6.2)	0.2% (-0.6, 1)	6.2% (4.0, 8.4)	1.6% (0.1, 3.1)	17.8% (15.3, 20.3)	4% (3.0, 5.0)	0% (-0.5, 0.5)

FIGURE 4.

All genotyped patients (N = 8140), adjusted probability of GT/ST by birth decade. Category (birth decade) percentages sum to 100. GT indicates genotype; ST, subtype.

FIGURE 5.

Living/uncured patients (N = 5138), unadjusted proportional distribution of select clinical factors across GT/ST. A, ALT category at most recent encounter, subset of living/uncured patients with available data (n = 3715). B, FIB4 category at most recent encounter, subset of living/uncured patients with available data (n = 3173). C, Charlson/Deyo comorbidity index as of most recent encounter, all living/uncured patients (N = 5138). ALT indicates alanine aminotransferase; FIB4, Fibrosis-4; GT, genotype; ST, subtype; ULN, upper limit of normal.



	Normal	ULN≤2xULN	>2xULN
GT1A	56.2% (54.5, 57.8)	30.9% (29.4, 32.3)	13.0% (11.8, 14.2)
GT1B	57.9% (55.0, 60.9)	28.4% (25.8, 31.1)	13.6% (11.6, 15.7)
GT1O	53.3% (47.6, 59.1)	33.0% (27.5, 38.5)	13.7% (9.4, 18.0)
GT2	63.6% (58.7, 68.5)	22.3% (17.8, 26.9)	14.1% (10.8, 17.4)
GT3	49.0% (43.6, 54.4)	29.9% (24.8, 35.0)	21.1% (17.3, 24.9)
GT4&6	57.9% (46.0, 69.9)	30.5% (19.7, 41.3)	11.6% (3.1, 20.0)

	≤1.21	1.21–5.88	>5.88
GT1A	25.5% (24.0, 27.0)	60.9% (59.2, 62.6)	13.6% (12.4, 14.8)
GT1B	20.7% (18.0, 23.5)	64.3% (61.3, 67.4)	14.9% (12.7, 17.2)
GT1O	21.1% (15.4, 26.7)	67.6% (61.3, 74.0)	11.3% (6.7, 15.9)
GT2	27.0% (22.7, 31.4)	64.8% (59.7, 69.8)	8.2% (4.5, 12.0)
GT3	25.2% (20.4, 30.0)	55.8% (50.3, 61.2)	19.0% (14.8, 23.2)
GT4&6	29.9% (19.1, 40.7)	44.6% (31.8, 57.4)	25.5% (16.7, 34.4)

FIGURE 6.

Living/uncured patients: adjusted probability of (N = 5138) ALT (A) and FIB4 (N = 3173) (B) by GT/ST. Genotype/subtype percentages sum to 100. ALT indicates alanine aminotransferase; FIB4, Fibrosis-4; GT, genotype; ST, subtype.

TABLE 1

All Genotyped Patients (N = 8140)

Variables	Response	Row Total (N = 8140) (Row %)	GT1 [n (%)]				GT2 [n (%)]				GT3 (N =827) (10.2%)	GT4 (N = 122) (1.5%)	GT6 (N = 26) (0.3%)	*P*
			All (N =6140) (75.4%)	1A (N =3596) (44.2%)	1B (N=1980) (24.3%)	Other (N =564) (9.2%)	All (N =1025) (12.6%)	2A (N =72) (0.9%)	2B (N=663) (8.1%)	Other (N =290) (3.6%)				
A: Distribution of HCV GT/ST across patient characteristics (row percentages shown for categories)														
Site	GHS	2076	1533 (74)	921 (44)	321 (15)	291 (14)	289 (14)	9 (0)	194 (9)	86 (4)	214 (10)	38 (2)	2 (0.1)	< 0.0001
	HFHS	3324	2781 (84)	1543 (46)	1120 (34)	118 (4)	214 (6)	50 (2)	135 (4)	29 (1)	270 (8)	55 (2)	4 (0.1)	
	KPHI	764	537 (70)	350 (46)	154 (20)	33 (4)	117 (15)	10 (1)	74 (10)	33 (4)	104 (14)	4 (1)	2 (0.3)	
	KPNW	1976	1289 (65)	782 (40)	385 (19)	122 (6)	405 (20)	3 (0)	260 (13)	142 (7)	239 (12)	25 (1)	18 (0.9)	
Sex	Male	4967	3792 (76)	2227 (45)	1224 (25)	341 (7)	589 (12)	36 (1)	389 (8)	164 (3)	511 (10)	58 (1)	17 (0.3)	0.0022
	Female	3173	2348 (74)	1369 (43)	756 (24)	223 (7)	436 (14)	36 (1)	274 (9)	126 (4)	316 (10)	64 (2)	9 (0.3)	
Race	Asian/Other	871	590 (68)	352 (40)	197 (23)	41 (5)	123 (14)	6 (1)	67 (8)	50 (6)	105 (12)	28 (3)	25 (2.9)	< 0.0001
	Black	1934	1812 (94)	932 (48)	789 (41)	91 (5)	87 (4)	26 (1)	44 (2)	17 (1)	25 (1)	10 (1)	0 (0)	
	White	5335	3738 (70)	2312 (43)	994 (19)	432 (8)	815 (15)	40 (1)	552 (10)	223 (4)	697 (13)	84 (2)	1 (0)	
Birth year	< 1946	814	643 (79)	272 (33)	331 (41)	40 (5)	128 (16)	21 (3)	74 (9)	33 (4)	26 (3)	12 (1)	5 (0.6)	< 0.0001
	1946≤1955	3692	2926 (79)	1628 (44)	1063 (29)	235 (6)	432 (12)	30 (1)	272 (7)	130 (4)	282 (8)	41 (1)	11 (0.3)	
	1956≤1965	2428	1737 (72)	1106 (46)	434 (18)	197 (8)	323 (13)	14 (1)	215 (9)	94 (4)	331 (14)	30 (1)	7 (0.3)	
	1966≤1975	586	410 (70)	265 (45)	100 (17)	45 (8)	81 (14)	6 (1)	55 (9)	20 (3)	78 (13)	14 (2)	3 (0.5)	
	≥1976	620	424 (68)	325 (52)	52 (8)	47 (8)	61 (10)	1 (0)	47 (8)	13 (2)	110 (18)	25 (4)	0 (0)	

Variables	Response	All Genotyped (N = 8140)	GT1 [n (%)]				GT2 [n (%)]				GT3 (N = 827)	GT4 (N = 122)	GT6 (N = 26)	*P**
			All (N = 6140)	1A (N = 3596)	1B (N = 1980)	Other (N = 564)	All (N = 1025)	2A (N = 72)	2B (N = 663)	Other (N = 290)				
B: Distribution of treatment and vital status by GT/ST (column percentages shown by category for each GT/ST)														
Treatment status	Untreated	4426 (54)	3448 (56)	2047 (57)	1119 (57)	282 (50)	503 (49)	42 (58)	328 (49)	133 (46)	400 (48)	61 (50)	14 (54)	< 0.0001
	Treated	3714 (46)	2692 (44)	1549 (43)	861 (43)	282 (50)	522 (51)	30 (42)	335 (51)	157 (54)	427 (52)	61 (50)	12 (46)	
	SVR	*1720 (21)*	*1055 (17)*	*579 (16)*	*357 (18)*	*119 (21)*	*371 (36)*	*15 (21)*	*248 (37)*	*108 (37)*	*254 (31)*	*30 (25)*	*10 (38)*	
	Treatment failure	*1994 (24)*	*1637 (27)*	*970 (27)*	*504 (25)*	*163 (29)*	*151 (15)*	*15 (21)*	*87 (13)*	*49 (17)*	*173 (21)*	*31 (25)*	*2 (8)*	
Treatment outcome (N = 3714)	SVR	1720 (46)	1055 (39)	579 (37)	357 (41)	119 (42)	371 (71)	15 (50)	248 (74)	108 (69)	254 (59)	30 (49)	10 (83)	

			GT1 [n (%)]				GT2 [n (%)]							
Variables	**Response**	**All Genotyped (N = 8140)**	**All (N = 6140)**	**1A (N = 3596)**	**1B (N = 1980)**	**Other (N = 564)**	**All (N = 1025)**	**2A (N = 72)**	**2B (N = 663)**	**Other (N = 290)**	**GT3 (N = 827)**	**GT4 (N = 122)**	**GT6 (N = 26)**	***P****
	Treatment failure	1994 (54)	1637 (61)	970 (63)	504 (59)	163 (58)	151 (29)	15 (50)	87 (26)	49 (31)	173 (41)	31 (51)	2 (17)	
Deceased	Yes	1486 (18)	1174 (19)	638 (18)	429 (22)	107 (19)	157 (15)	18 (25)	99 (15)	40 (14)	137 (17)	14 (11)	4 (15)	0.0276
	No	6654 (82)	4966 (81)	2958 (82)	1551 (78)	457 (81)	868 (85)	54 (75)	564 (85)	250 (86)	690 (83)	108 (89)	22 (85)	

*P-value for genotype based on χ^2 test; subtype should be considered exploratory.

GHS indicates Geisinger Health System; GT, genotype; HFHS, Henry Ford Health System; KPHI, Kaiser-Permanente Hawai'I; KPNW, Kaiser-Permanente Northwest; SVR, sustained virological response; ST, subtype.

TABLE 2

Living/Uncured Patients (N = 5138)

Variables	Response	Row Total (N = 5138) (Row %)	GT1 [n (%)] All (N = 4056) (78.9%)	1A (N =2470) (48.1%)	1B (N=1236) (24.1%)	Other (N = 350) (6.9%)	GT2 [n (%)] All (N = 528) (10.3%)	2A (N=40) (0.8%)	2B (N=338) (6.6%)	Other (N=150) (2.9%)	GT3 (N = 461) (9.0%)	GT4 (N = 81) (1.6%)	GT6 (N = 12) (0.2%)	*P*
A: Distribution of HCV GT/ST across patient characteristics (row percentages shown for categories)														
Site	GHS	1257	962 (77)	621 (49)	170 (14)	171 (14)	150 (12)	7 (1)	105 (8)	38 (3)	118 (9)	27 (2)	0 (0)	
	HFHS	2013	1749 (87)	990 (49)	690 (34)	69 (3)	94 (5)	25 (1)	55 (3)	14 (1)	138 (7)	32 (2)	1 (0)	
	KPHI	521	392 (75)	263 (50)	104 (20)	25 (5)	65 (12)	6 (1)	41 (8)	18 (3)	60 (12)	2 (0)	2 (0)	
	KPNW	1346	953 (71)	596 (44)	272 (20)	85 (6)	219 (16)	2 (0)	137 (10)	80 (6)	145 (11)	20 (1)	9 (1)	
Sex	Male	3058	2431 (79)	1475 (48)	750 (25)	206 (7)	300 (10)	15 (0)	198 (6)	87 (3)	284 (9)	36 (1)	7 (0)	
	Female	2079	1625 (78)	995 (48)	486 (23)	144 (7)	228 (11)	25 (1)	140 (7)	63 (3)	177 (9)	45 (2)	5 (0)	
Race	Asian/Other	549	393 (72)	238 (43)	126 (23)	29 (5)	70 (13)	5 (1)	36 (7)	29 (5)	58 (11)	16 (3)	12 (2)	
	Black	1322	1253 (95)	666 (50)	526 (40)	61 (5)	48 (4)	15 (1)	23 (2)	10 (1)	15 (1)	7 (1)	0 (0)	
	White	3266	2410 (74)	1566 (48)	584 (18)	260 (8)	410 (13)	20 (1)	279 (9)	111 (3)	388 (12)	58 (2)	0 (0)	
Birth year	< 1946	453	368 (81)	162 (36)	182 (40)	24 (5)	63 (14)	10 (2)	41 (9)	12 (3)	11 (2)	8 (2)	3 (1)	
	1946≤1955	2226	1862 (84)	1073 (48)	651 (29)	138 (6)	195 (9)	14 (1)	121 (5)	60 (3)	136 (6)	30 (1)	3 (0)	
	1956≤1965	1590	1189 (75)	771 (48)	295 (19)	123 (8)	180 (11)	10 (1)	114 (7)	56 (4)	198 (12)	19 (1)	4 (0)	
	1966≤1975	411	301 (73)	196 (48)	70 (17)	35 (9)	51 (12)	5 (1)	32 (8)	14 (3)	52 (13)	6 (1)	2 (0)	
	≥1976	457	336 (74)	268 (59)	38 (8)	30 (7)	39 (9)	1 (0)	30 (7)	8 (2)	64 (14)	18 (4)	0 (0)	

Variables	Response	Living/ uncured (N = 5138)	GT2 [n (%)] All (N = 4056)	1A (N = 2470)	1B (N = 1236)	Other (N = 350)	GT1 [n (%)] All (N = 528)	2A (N = 40)	2B (N = 338)	Other (N = 150)	GT3 (N = 461)	GT4 (N = 81)	GT6 (N = 12)	P-value (for "All" GT only)
B: Distribution of clinical characteristics by GT/ST (column percentages shown by category for each GT/ST)														
Treatment status	Treatment failure	1721 (33)	1412 (35)	855 (35)	422 (34)	135 (39)	132 (25)	11 (28)	81 (24)	40 (27)	146 (32)	30 (37)	1 (8)	< 0.0001
	Untreated	3417 (66)	2644 (65)	1615 (65)	814 (66)	215 (61)	396 (75)	29 (73)	257 (76)	110 (73)	315 (68)	51 (63)	11 (92)	
ALT available (n = 3715)	<LLN and/or Normal	2101 (6)	1694 (57)	1008 (56)	565 (62)	121 (49)	233 (61)	21 (78)	150 (60)	62 (59)	137 (44)	31 (54)	7 (58)	< 0.0001
	ULN≤2xULN	1092 (3)	883 (30)	557 (31)	239 (26)	87 (35)	90 (23)	3 (11)	61 (24)	26 (25)	97 (31)	19 (33)	3 (25)	
	> 2xULN	522 (1)	378 (13)	235 (13)	104 (11)	39 (16)	60 (16)	3 (11)	40 (16)	17 (16)	76 (25)	7 (12)	1 (8)	
FIB4 available (n = 3173)	≤1.21	772 (24)	596 (24)	405 (26)	136 (18)	55 (25)	87 (27)	11 (50)	59 (29)	17 (18)	72 (26)	14 (29)	3 (25)	< 0.0001
	1.21≤5.88	1961 (62)	1584 (63)	942 (61)	502 (67)	140 (64)	206 (64)	10 (45)	125 (61)	71 (76)	147 (54)	21 (43)	3 (25)	

Variables	Response	Living/ uncured (N = 5138)	GT2 [n (%)]				GT1 [n (%)]				GT3 (N = 461)	GT4 (N = 81)	GT6 (N = 12)	P-value (for "All" GT only)
			All (N = 4056)	1A (N = 2470)	1B (N = 1236)	Other (N = 350)	All (N = 528)	2A (N = 40)	2B (N = 338)	Other (N = 150)				
	> 5.88	440 (14)	340 (13)	207 (13)	109 (15)	24 (11)	29 (9)	1 (5)	22 (11)	6 (6)	54 (20)	14 (29)	3 (25)	
Decompensated cirrhosis	Yes	134 (3)	106 (3)	53 (2)	49 (4)	4 (1)	11 (2)	2 (5)	8 (2)	1 (1)	16 (3)	1 (1)	0 (0)	0.288
	No	5004 (97)	3950 (97)	2417 (98)	1187 (96)	346 (99)	517 (98)	38 (95)	330 (98)	149 (99)	445 (97)	80 (99)	12 (100)	
Weighted Charlson/Deyo Index	0	3252 (63)	2552 (63)	1602 (65)	712 (58)	238 (68)	328 (62)	22 (55)	209 (62)	97 (65)	319 (69)	46 (57)	7 (58)	0.001
	1	797 (16)	626 (15)	378 (15)	199 (16)	49 (14)	101 (19)	8 (20)	66 (20)	27 (18)	57 (12)	10 (12)	3 (25)	
	2	394 (8)	301 (7)	185 (7)	94 (8)	22 (6)	43 (8)	5 (13)	26 (8)	12 (8)	41 (9)	8 (10)	1 (8)	
	3	695 (14)	577 (14)	305 (12)	231 (19)	41 (12)	56 (11)	5 (13)	37 (11)	14 (9)	44 (10)	17 (21)	1 (8)	
Specific comorbidities (nonexclusive)	Diabetes	381 (7)	315 (8)	179 (7)	112 (9)	24 (7)	32 (6)	1 (3)	26 (8)	5 (3)	19 (4)	13 (16)	2 (17)	0.0002
	Alcohol abuse	513 (10)	378 (9)	256 (10)	97 (8)	25 (7)	55 (10)	3 (8)	37 (11)	15 (10)	73 (16)	7 (9)	0 (0)	< 0.0001
	HBV coinfection	46 (1)	24 (1)	9 (0)	14 (1)	1 (0)	10 (2)	0 (0)	7 (2)	3 (2)	10 (2)	1 (1)	1 (8)	< 0.0001
	HIV coinfection	158 (3)	127 (3)	80 (3)	38 (3)	9 (3)	16 (3)	1 (3)	10 (3)	5 (3)	11 (2)	4 (5)	0 (0)	0.2881

*P-value for genotype based on χ^2 test; subtype should be considered exploratory.

ALT indicates alanine aminotransferase; FIB4, Fibrosis-4; GHS, Geisinger Health System; GT, genotype; HFHS, Henry Ford Health System; KPHI, Kaiser-Permanente Hawai'I; KPNW, Kaiser-Permanente Northwest; ST, subtype.